Encompassing test for parametric and nonparametric regression techniques

Patrick Rakotomarolahy

Encompassing test for parametric and nonparametric regression techniques

Patrick Rakotomarolahy*

Abstract

This paper examines encompassing test for parametric and nonparametric methods. We provide the asymptotic normality of the encompassing statistic associated to the encompassing hypothesis with parametric and nonparametric regression methods. We develop various results on this test for more general processes satisfying several dependence structures. Moreover, we apply the encompassing test on real economic activity modelling.

Keywords: encompassing test, functional parameter, mixing processes, nonparametric techniques, asymptotic normality.
JEL: C22 - C53 - E32.

1 Introduction

Model selection is a challenging step in statistical modelling of economics and finance. Modelling economic or financial data requires characterization of the associated data generating process (DGP). The DGP is unknown and therefore we face several admissible competing models. Model selection consists on selecting a model, which mimics such unknown DGP, from a set of admissible models according to a criterion. One retains the model which makes such criterion optimal. There exist various model selection criteria in the literature when admissible models have fully parametric specification, such as the Wald test, the likelihood ratio test, the Lagrange multiplier test, the information criteria and so on. The other case, that is when admissible models contain simultaneously parametric and nonparametric specifications, seems underdeveloped.

*Centre d'Economie de la Sorbonne, Université Paris 1 Panthéon-Sorbonne, 106 boulevard de l'Hopital 75647 Paris Cedex 13, France, e-mail: rakotopapa@yahoo.fr.

Encompassing test appears to be helpful for the latter situation where an encompassing model is intended to account the salient feature of the encompassed model. Therefore, encompassing test can detect redundant models among the admissible models.

Encompassing tests are based on two points, that the encompassing model ought to be able to explain the predictions and to predict some mis-specifications of the encompassed model, Hendry et al. (2008). We know that there are various considerations and developments of encompassing tests, we refer readers to Mizon (1984), Hendry and Richard (1989), Gouriéroux and Monfort (1995) and Florens *et al.* (1996). For an overview on the concept of the encompassing test, see Bontemps and Mizon (2008) and Mizon (2008). Applications of encompassing tests can be found inside the model selection procedure of GETS modelling developed by Hendry and Doornik (1994) and Hoover and Perez (1999).

Recently, Bontemps et al. (2008) have developed encompassing tests which cover large set of methods such as parametric and nonparametric methods. Among their results, encompassing tests with nonparametric methods are established, considering kernel regression method. They provide asymptotic normality of the associated encompassing statistics under the independent and identically distributed hypothesis (i.i.d).

We extend the results of Bontemps et al. (2008) in two directions. First, we extend their encompassing tests using the nearest neighbor regression method which has be claimed more flexible compared to kernel regression method. Other motivation on developing encompassing test for nearest neighbor regression would be its consideration in the literature for different applications in finance as well as in economics for capturing nonlinear features of the financial and economic data sets, Mizrach (1992), Nowman and Saltoglu (2003), Guégan and Huck (2005) and Guégan and Rakotomarolahy (2010), among others. In independent framework, we achieve similar asymptotic normality result as in Bontemps et al. (2008) for the encompassing test under some regularity conditions.

Second, we relax the independent hypothesis by focusing on processes with some dependence structures. This second extension lies on the generalization of encompassing test for dependent processes.

We will provide an application on real economic activity modelling. We will adress variable selection as well as model selection problems when modelling the Gross Domestic Product.

The paper is organized as follows. After a brief introduction for the motivation on developing this encompassing test, we will make an overview of such test in section 2. In section 3, we will provide new results on the asymptotic normality of the encompassing test associated to linear parametric modelling and nonparametric kernel and nearest neighbor regression methods. In section 4, we will make an illustration of the results on real data and last we conclude.

2 Preliminary study

This section introduces the encompassing test and then builds the corresponding encompassing hypothesis. So, given two regression models $\mathcal{M}_1$ and $\mathcal{M}_2$, we are interested in knowing if the model $\mathcal{M}_1$ can account the result of model $\mathcal{M}_2$. In other words, we want to know if $\mathcal{M}_1$ encompasses $\mathcal{M}_2$ or in a short notation $\mathcal{M}_1 \mathcal{E} \mathcal{M}_2$. Testing such a hypothesis will be done using the notion of encompassing test.

Generally speaking, model $\mathcal{M}_1$ encompasses model $\mathcal{M}_2$, if the parameter $\theta_{\mathcal{M}_2}$ of the latter model can be expressed in function of the parameter $\theta_{\mathcal{M}_1}$ of the former model. In other words, let $\Delta(\theta_{\mathcal{M}_1})$ be the pseudo true value of $\theta_{\mathcal{M}_2}$ on $\mathcal{M}_1$. In general, the pseudo-true value is defined as the plim of $\hat{\theta}_{\mathcal{M}_2}$ on $\mathcal{M}_1$, Bontemps et al. (2008). For more discussion on pseudo-true value associated with the KLIC[1], we refer to Sawa (1978) and Govaerts et al. (1994) among others. The encompassing statistic is given by the difference between $\hat{\theta}_{\mathcal{M}_2}$ and $\Delta(\hat{\theta}_{\mathcal{M}_1})$ scaled by a coefficient a_n. Specification of the encompassing test will depend on the estimation of the regression method: parametric or nonparametric methods.

Let $S = (Y, X, Z)$ be a zero mean random process with valued in $\mathbb{R}$x$\mathbb{R}^d$x$\mathbb{R}^q$ where $d, q \in \mathbb{N}^*$. For $x \in \mathbb{R}^d$ and $z \in \mathbb{R}^q$, consider the two models $\mathcal{M}_1$ and $\mathcal{M}_2$ as the conditional expectations $m(x)$

[1] Kullback-Leiber information criterion

and $g(z)$, respectively and are defined as follows:

$$\mathcal{M}_1 : \quad m(x) = E[Y|X = x] \quad and \quad \mathcal{M}_2 : \quad g(z) = E[Y|Z = z] \tag{2.1}$$

Moreover, the general unrestricted model is given by $r(x, z) = E[Y|X = x, Z = z]$. We follow the encompassing test in Bontemps *et al.* (2008).
We are interested in testing the hypothesis that $\mathcal{M}_1$ encompasses $\mathcal{M}_2$, and then introducing the null hypothesis:

$$\mathcal{H} : E[Y|X = x, Z = z] = E[Y|X = x]. \tag{2.2}$$

This null states that $\mathcal{M}_1$ is the owner model, and $\mathcal{M}_2$ will be served on validating this statement and is called the rival model. We test this hypothesis $\mathcal{H}$ through the following implicit encompassing hypothesis:

$$\mathcal{H}^* : E[E[Y|X = x]/Z = z] = E[Y|Z = z]. \tag{2.3}$$

The following homoskedasticity condition will be assumed all along this chapter:

$$Var[Y|X = x, Z = z] = \sigma^2. \tag{2.4}$$

Moreover, a necessary condition for the encompassing test relies on the errors of both models where the intended encompassing model $\mathcal{M}_1$ should have smaller standard error than the encompassed model $\mathcal{M}_2$.

In general, $\mathcal{M}_1$ or $\mathcal{M}_1$ can be estimated using nonparametric or parametric regression method. We will consider these different situations when the processes $(S_n)_n$ are independent or dependent. We begin by constructing the encompassing statistic associated to each of these four situations and after we discuss their asymptotic behaviors.

3 Encompassing statistic

3.1 General framework

We are interested on the asymptotic behavior of the encompassing statistic associated to the null hypothesis $\mathcal{M}_1\mathcal{E}\mathcal{M}_2$. We can encounter the following four situations: $\mathcal{M}_1$ and $\mathcal{M}_2$ are both estimated parametrically, $\mathcal{M}_1$ and $\mathcal{M}_2$ are both estimated nonparametrically, $\mathcal{M}_1$ is estimated nonparametrically and $\mathcal{M}_2$ parametrically and $\mathcal{M}_1$ is estimated parametrically and $\mathcal{M}_2$ nonparametrically. We will consider the kernel and the k-NN regression estimates for nonparametric methods and the linear regression for the parametric methods. For both independent and dependent processes, we will study and establish the asymptotic normality of the corresponding four encompassing tests.

Consider a sample $S_i = (Y_i, X_i, Z_i)$, $i = 1, ..., n$, which can be viewed as realization of the random process $S = (Y, X, Z)$ with valued in $\mathbb{R}\text{x}\mathbb{R}^d\text{x}\mathbb{R}^q$ where $d, q \in \mathbb{N}^*$. We suppose that S_i, $i = 1, ..., n$ has a joint density f. Moreover, $\varphi(.,.)$, $\varphi(. \mid .)$ and $\varphi(.)$ will denote the joint, the conditional and the marginal densities of the process (Y, Z), respectively. That is, for $y \in \mathbb{R}$ and $z \in \mathbb{R}^q$, $\varphi(y, z)$, $\varphi(y \mid z)$ and $\varphi(z)$ correspond to the density of the following processes (Y, Z) at point (y, z), $(Y \mid Z = z)$ at point y and Z at point z, respectively. Similarly, h will denote the joint, the conditional and the marginal densities of the process (Y, X), according to the argument that it takes.

Throughout this section, we assume the existence of continuous version of the various joint and marginal density functions and of the three conditional means m, g and r. In addition, the square integrability will be assumed.

In this paragraph, $N(\mu, v)$ will denote the Gaussian distribution with mean μ and variance v. We now consider the first case that is the encompassing test when the two models $\mathcal{M}_1$ and $\mathcal{M}_2$ have parametric specification.

3.2 Parametric modelling vs parametric modelling

Encompassing test for parametric modelling has been developed a lot in the literature. We discuss briefly one parametric encompassing test where models $\mathcal{M}_1$ and $\mathcal{M}_2$ have linear parametric specification. In that case, the two models $\mathcal{M}_1$ and $\mathcal{M}_2$ are given by (3.1) and the nesting model is represented by the function r:

$$\begin{aligned} &m(x) = \beta' x \ \ with \ \ \beta = (E[XX'])^{-1} E[XY] \\ &g(z) = \gamma' z \ \ with \ \ \gamma = (E[ZZ'])^{-1} E[ZY] \\ &r(x{,}z) = \alpha' w \ \ with \ \ \alpha = E[WW']^{-1} E[WY] \ \ and \ \ W = (X, Z). \end{aligned} \tag{3.1}$$

We can get the estimates $\hat{\beta}$, $\hat{\gamma}$ and $\hat{\alpha}$ of the paramaters β, γ and α, respectively, using the sample $S_i = (Y_i, X_i, Z_i)$, $i = 1, ..., n$. Now, testing $\mathcal{M}_1 \mathcal{E} \mathcal{M}_2$ corresponds to the test of the null hypothesis $\mathcal{H}$ where the conditional mean is just the linear projection. Therefore, the encompassing statistic of the null $\mathcal{M}_1 \mathcal{E} \mathcal{M}_2$ can be written as follows.

$$\hat{\delta}_{\beta,\gamma} = \hat{\gamma} - \hat{\gamma_L}(\hat{\beta}), \tag{3.2}$$

where $\hat{\gamma_L}(\hat{\beta})$ is an estimate of the pseudo-true value $\gamma_L(\beta)$ associated with $\hat{\gamma}$ on $\mathcal{H}_1$. Remarking that the pseudo-true value is defined by $\gamma_L(\beta) = (E[ZZ'])^{-1}E[ZX']\beta$, we state in the following theorem the asymptotic behavior of the encompassing statistic in relation (3.2).

Theorem 3.1. *Assume that the relation 2.4 is satisfied. When the sample $S_i = (Y_i, X_i, Z_i)$, $i = 1, ..., n$ are i.i.d., then under $\mathcal{H}$, we get:*

$$\sqrt{n}\hat{\delta}_{\beta,\gamma} \to N(0, \sigma^2\Omega) \quad \textit{in} \quad \textit{distribution} \quad \textit{as} \quad n \to \infty. \tag{3.3}$$

where $\Omega = Var(Z)^{-1}E[Var(Z \mid X)]Var(Z)^{-1}$.

For development on this asymptotic behavior of the encompassing statistic, we refer to Gouriéroux et al. (1983) and Mizon and Richard (1986) among others. For recent discussion on this encompassing test for fully parametric case, Bontemps et al. (2008) is a good reference.

Development of the parametric encompassing test goes beyond independent processes in the literature. We may mention the remark in Bontemps et al (2008) stating the obvious extension of the asymptotic results for independent processes to stationary ergodic processes in line with White (1990a). Moreover, encompassing test for dynamic stationary models and time series regressions have been discussed in Govaerts et al. (1994), Hendry (1995), Hendry and Nielsen (2006), among others.

Next, we will study the completely nonparametric case.

3.3 Nonparametric modelling for $\mathcal{M}_1$ and $\mathcal{M}_2$

We now consider the case where the two models $\mathcal{M}_1$ and $\mathcal{M}_2$ defined in (2.1) are estimated using nonparametric techniques. To test the hypothesis " $\mathcal{M}_1$ encompasses $\mathcal{M}_2$ ", we build the corresponding encompassing statistic and establish asymptotic property of such statistic.

Using the sample $S_i = (Y_i, X_i, Z_i)$, $i = 1, ..., n$ and the associated functional estimates m_n and g_n of the unknown functions m and g in relation (2.1) respectively, we define the encompassing statistic as follows:

$$\hat{\delta}_{m,g}(z) = g_n(z) - \hat{G}(m_n)(z), \tag{3.4}$$

where $\hat{G}(m_n)$ is an estimate of the pseudo true value $G(m)$ associated with g_n on $\mathcal{H}$, which is defined by $G(m) = E[m \mid Z = z]$. We focus on the following two nonparametric regression function estimates: the kernel and the k-NN regression function estimates.

a) Encompassing test when m and g are estimated using kernel regression estimate

We estimate the unknown conditional means m and g defined in (2.1) using kernel regression estimate. We consider the sample $S_i = (Y_i, X_i, Z_i)$, $i = 1, ..., n$, which is a realization of the random process $S = (Y, X, Z)$. Then, the kernel regression estimates m_n of the function m, and g_n of the function g have the following expressions:

$$m_n(x) = \frac{\frac{1}{nh_{1n}^d}\sum_{i=1}^n K_1(\frac{x-X_i}{h_{1n}})Y_i}{\frac{1}{nh_{1n}^d}\sum_{i=1}^n K_1(\frac{x-X_i}{h_{1n}})} \quad g_n(z) = \frac{\frac{1}{nh_{2n}^q}\sum_{i=1}^n K_2(\frac{z-Z_i}{h_{2n}})Y_i}{\frac{1}{nh_{2n}^q}\sum_{i=1}^n K_2(\frac{z-Z_i}{h_{2n}})} \tag{3.5}$$

where h_{jn} and K_j, $j = 1, 2$ are window widths and kernel densities, respectively. The kernel densities satisfy

$$K_j(u) \geq 0 \ \ and \ \ \int K_j(u)du = 1 \ \ j = 1, 2. \tag{3.6}$$

We separate the independent case and the dependent case.

a1) Independent case

We establish asymptotic normality of the encompassing statistic defined in equation (3.4). Bontemps et al. (2008) has provided asymptotic property of encompassing statistic for the kernel regression function estimate under the independent hypothesis and some assumptions. In addition to the usual kernel regression condition, we need the following assumption to eliminate the bias.

Assumption 3.1. *Assume that* $\lim_{n\to\infty} n.h_{1n}^{d+2p} = 0$ *and* $\lim_{n\to\infty} n.h_{2n}^{q+2p} = 0$ *with* p *is the order of differentiability of the density and regression functions.*

Under the previous assumption, the bias of the kernel regression disappears, for discussion we refer to Vieu (1994) and Bontemps et al. (2008), among others.
The last assumption concerns the speed of convergence of the smoothing parameters.

Assumption 3.2. *Assume that the window widths satisfy:* $\lim_{n\to\infty} n.h_{1n}^{d} = 0$, $\lim_{n\to\infty} n.h_{2n}^{q} = 0$ *and* $\lim_{n\to\infty} \frac{h_{2n}^q}{h_{1n}^d} = 0$.

Now, we can establish the asymptotic normality of the encompassing statistic.

Theorem 3.2. *Assume that relation (2.4) is satisfied. When the kernel densities* K_j *and the window widths* h_{jn}, $j = 1, 2$ *satisfy the kernel regularity condition, given in assumptions 3.1 and 3.2, under* $\mathcal{H}$, *we get:*

$$\sqrt{nh_{2n}^q}\,\hat{\delta}_{m,g}(z) \to N\Big(0, \frac{\sigma^2 \int K_2^2(u)du}{\varphi(z)}\Big) \quad \text{in distribution as} \quad n \to \infty. \tag{3.7}$$

$\varphi(z)$ *is the marginal density of the* Z *at* z.

For the proof of this theorem, we refer to Bontemps et al. (2008). We now move to the case that the processes exhibit some dependence structures.

a2) Dependent case

For dependent processes, to get the asymptotic normality of the associated encompassing statistic, we need the following assumptions. The assumptions have been drawn from Bosq (1998). The first assumption characterizes the dependence structure.

Assumption 3.3. (S_t) *is* α*-mixing with* $\alpha(n) = O(n^{-\rho})$ *where* $\rho > \frac{\nu^2+4}{2\nu}$ *for some positive* ν.

The next assumption collects regularity conditions on the continuity and on the differentiability of the density functions.

Assumption 3.4. φ *and* $g*\varphi$ *are* $C_{2,d}(b)$ *for some real* b *where* $C_{2,d}(b)$ *the space of twice continuously differentiable real valued functions* f, *defined on* $\mathbb{R}^d$, *such that* $||\varphi||_\infty \leq b$ *and* $||\varphi^{(2)}||_\infty \leq b$ *with* $\varphi^{(2)}$ *denotes any partial derivative of order 2 for* φ. *Next,* $Sup_{t\geq k}||\varphi(Z_1, Z_t)||_\infty < \infty$ *and last,* $\varphi(.)E[Y_1^2|Z_1 = .]$ *is continuous at* z.

The last assumption concerns finiteness on the moments of the process $(Y_n, Z_n)_n$.

Assumption 3.5. *$||E[|Y_1|^{4+\nu}|Z_1 = .]||_\infty < \infty$; $E[|Z_1|^{4+\nu}] < \infty$ for some positive ν; $Sup_{t\in\mathbb{N}}||E[Y_1^i Y_t^j|Z_t = ., Z_1 = .]||_\infty < \infty$ where $i \geq 0$, $j \geq 0$, $i+j = 2$.*

We provide in the following, a theorem establishing the asymptotic convergence of the encompassing statistic.

Theorem 3.3. *Suppose that assumptions 3.3-3.5 hold. Moreover, suppose that relation (3.11) is satisfied. Then, under $\mathcal{H}$, we get:*

$$\sqrt{nh_{2n}^q}\hat{\delta}_{m,g}(z) \to N(0, \frac{\sigma^2 \int K_2^2(u)du}{\varphi(z)}) \quad in \quad distribution \quad as \quad n \to \infty. \tag{3.8}$$

$\varphi(z)$ is the marginal density of the Z at z and $\sigma^2 = Var[Y|X = x, Z = z]$.

Proof of theorem 3.3 The proof of this theorem will be based on the decomposition of the expression of the encompassing statistic into two parts as follows:

$$\begin{aligned}
\sqrt{nh_{2n}^q}\hat{\delta}_{m,g}(z) &= \sqrt{nh_{2n}^q}(g_n(z) - \hat{G}(m_n)(z)) \\
&= \sqrt{nh_{2n}^q}(\sum_{t=1}^{n} \frac{K_2(\frac{z-Z_t}{h_{2n}})}{\sum_{t=1}^n K_2(\frac{z-Z_t}{h_{2n}})} Y_t - \sum_{t=1}^{n} \frac{K(\frac{z-Z_t}{h_n})}{\sum_{t=1}^n K(\frac{z-Z_t}{h_n})} m_n(x_t)) \\
&= \sqrt{nh_{2n}^q}\sum_{t=1}^{n} \frac{K_2(\frac{z-Z_t}{h_{2n}})}{\sum_{t=1}^n K_2(\frac{z-Z_t}{h_{2n}})}(Y_t - m(x_t)) \qquad (3.9) \\
&+ \sqrt{nh_{2n}^q}\sum_{t=1}^{n} \frac{K(\frac{z-Z_t}{h_n})}{\sum_{t=1}^n K(\frac{z-Z_t}{h_n})}(m(x_t) - m_n(x_t)) \\
&= C_1 + C_2.
\end{aligned}$$

The first part C_1 coincides to the kernel regression of the residuals $\epsilon_t = Y_t - m(x_t)$ onto Z_t. When assumptions 3.3-3.5 hold, then under $\mathcal{H}$, we achieved the convergence in distribution of the first part to a normal distribution using Rhomari's result in Bosq (1998). The second part C_2 reflects the limit in probability of the supremum of the difference $m_n(x_t) - m(x_t)$ at $x_t \in \mathbb{R}^d$ scaled by $\sqrt{nh_n^q}$ and its convergence can be derived from the rate of covergence of the uniform convergence of the estimate $m_n(x_t)$ which has been provided by Bosq (1998).

We are interested on establishing similar asymptotic results for encompassing test associated to other nonparametric methods. So, instead of considering the kernel regression, we now consider the nearest neighbor regression.

b) Encompassing test when m and g are estimated using k-NN regression estimate

The previous result can be extended to the nonparametric nearest neighbor regression estimate with different assumptions. Then, we establish asymptotic distribution of the encompassing statistic $\hat{\delta}_{m,g}$ in relation (3.4) when the conditional mean functions m and g defined in (2.1) are estimated using the k-NN regression method. For example the k-NN regression estimate g_n of the conditional mean g can be written as follows:

$$g_n(z) = \frac{\frac{1}{nR_n^q}\sum_{i=1}^n w(\frac{z-Z_i}{R_n})Y_i}{\frac{1}{nR_n^q}\sum_{i=1}^n w(\frac{z-Z_i}{R_n})} \tag{3.10}$$

where R_n will be defined as distance, according to the Euclidean norm in $\mathbb{R}^q$, from z to its k-th neighbors with k the number of neighbors, and $w(u)$ is a bounded, non negative function satisfying

$$\int w(u)du = 1 \quad and \quad w(u) = 0 \quad for \quad |u| \geq 1 \tag{3.11}$$

To get the asymptotic normality of encompassing test, we need some assumptions that we state now.

b1) Independent case

When we work with independent processes, we will use the following assumptions which are the same as assumptions introduced in Mack(1981).

The first assumption relies on the density function φ of the process (Y, Z).

Assumption 3.6. *The function $\chi_\beta(z) = \int y^\beta \varphi(z, y)dy$ is bounded and continuous at z for $\beta = 0, 1, 2,$ and continuously differentiable in a neighborhood of z for $\beta = 0, 1$.*

The next assumption concerns conditions on the moments of Y up to order three.

Assumption 3.7. *$E|Y|^3 < \infty$, $Var(Y \mid Z = z) > 0$ and $f(z) > 0$.*

The last assumption states conditions on the relationship between the number of neighbors k and the sample size n.

Assumption 3.8. *$k = o(n)$, $log(n) = o(k)$ and $k \to \infty$ as $n \to \infty$.*

When assumptions 3.6-3.8 hold and the relation (3.11) is satisfied, then Mack (1981) has established the asymptotic normality of the centered k-NN regression of g_n. Moreover, we need also that the bias of such k-NN regression estimate vanishes to zero. To guarantee the latter statement, we complete the three previous assumptions with the following one:

Assumption 3.9. *$k = n^{\alpha}$ with $0 < \alpha < \frac{4}{4+d}$.*

We now provide the asymptotic normality of the encompassing statistic $\hat{\delta}_{m,g}$ when m and g are estimated using the k-NN regression method.

Theorem 3.4. *Assume that relation (2.4) is satisfied. When assumptions 3.6-3.9 and the relation (3.11) are satisfied, under $\mathcal{H}$, we have:*

$$\sqrt{k-1}\hat{\delta}_{m,g}(z) \to N(0, c.\sigma^2 \int w^2(u)du) \quad in \quad distribution \quad as \quad n \to \infty. \tag{3.12}$$

where $c = \frac{\pi^{q/2}}{\Gamma((q+2)/2)}$ is the volume of unit ball in $\mathbb{R}^q$ with $\Gamma(.)$ the gamma function.

Proof of Theorem 3.4. We mention that the functional parameters m_n and g_n are from k-NN regression estimate. We introduce the following notation for the weighting function in relation (3.10):

$$W(\frac{z-Z_i}{R_n}) = \frac{\frac{1}{nR_n^q} w(\frac{z-Z_i}{R_n})}{\frac{1}{nR_n^q}\sum_{i=1}^n w(\frac{z-Z_i}{R_n})} \tag{3.13}$$

The estimate $\hat{G}(m_n)$ of the pseudo true value $G(m) = E[m \mid Z]$ associated with g_n on $\mathcal{H}$ at a given point z is given by:

$$\hat{G}(m_n)(z) = \sum_{i=1}^{n} W_1(\frac{z-Z_i}{R_{1n}}) m_n(x_i) \tag{3.14}$$

where W_1 is a weighting function which is defined as follows:

$$W_1(\frac{z-Z_i}{R_{1n}}) = \frac{\frac{1}{nR_{1n}^q} w_1(\frac{z-Z_i}{R_{1n}})}{\frac{1}{nR_{1n}^q}\sum_{i=1}^n w_1(\frac{z-Z_i}{R_{1n}})} \tag{3.15}$$

where the function w_1 satisfies condition in relation (3.11). We now write down our encompassing statistic by replacing the estimates g_n and $\hat{G}(m_n)$ as follows:

$$\begin{aligned}\sqrt{k-1}\hat{\delta}_{m,g}(z) &= \sqrt{k-1}(g_n(z) - \hat{G}(m_n)(z)) \\ &= \sqrt{k-1}\sum_{i=1}^{n} W(\frac{z-Z_i}{R_n})Y_i - \sqrt{k-1}\sum_{i=1}^{n} W_1(\frac{z-Z_i}{R_{1n}}) m_n(x_i)\end{aligned} \tag{3.16}$$

Following similar techniques as in Bontemps *et al.* (2008), we decompose this statistic into three parts:

$$\begin{aligned}\sqrt{k-1}\hat{\delta}_{m,g}(z) &= \sqrt{k-1}\sum_{i=1}^{n} W(\frac{z-Z_i}{R_n})(Y_i - m(x_i)) + \sqrt{k-1}\sum_{i=1}^{n} W_1(\frac{z-Z_i}{R_{1n}})(m(x_i) - m_n(x_i)) \\ &= A_1 + A_2.\end{aligned} \tag{3.17}$$

This equality (3.17) is obvious when one considers the same weighting functions for both k-NN estimates g_n $G(m_n)$, that is $W = W_1$. Otherwise, we have extra term $\sqrt{k-1}(\sum_{i=1}^{n} W(\frac{z-Z_i}{R_n})m(x_i) - \sum_{i=1}^{n} W_1(\frac{z-Z_i}{R_{1n}})m(x_i))$, which can be viewed as difference bewteen two k-NN regression estimates of $m(x_i)$. We know that some regularity conditions k-NN regression estimate is asymptotically unbiased. Then the equality vanishes to zero. Thus, the equality (3.17) holds. We now proceed on analysing the two expressions A_1 and A_2 of this equality (3.17).

For A_1 which is the first expression in RHS of the first equality in relation (3.17), it involves a k-NN regression of $\epsilon_i = Y_i - m(x_i)$ on Z_i scaled by the coefficient $\sqrt{k-1}$ which indicates the convergence speed rate when n goes to infinity. From the asymptotic normality result due to Mack (1981), when assumptions 3.6 - 3.9 and relation (3.11) are satisfied, under $\mathcal{H}$ we have:

$$A_1 \to N(0, c.Var(\epsilon/Z = z)\int w^2(u)du) \quad in \quad distribution \quad as \quad n \to \infty. \tag{3.18}$$

Next, for the second expression A_2, we can bound by taking its supremum with respect to x_i and then we get:

$$\begin{aligned}|A_2| &\leq Sup_{x_i}\sqrt{k-1}|m_n(x_i) - m(x_i)| \\ &\leq Sup_{x_i}\sqrt{k-1}|m_n(x_i) - E[m_n(x_i)]| + Sup_{x_i}\sqrt{k-1}|E[m_n(x_i)] - m(x_i)| \\ &= B1 + B2\end{aligned} \tag{3.19}$$

When using the expression of the bias, theorem 1 in Mack(1981), B_2 becomes:

$$B_2 = (Sup_{x_i}A(x_i))(\frac{k}{n})^{\frac{2}{d}}\sqrt{k-1} + o((\frac{k}{n})^{\frac{2}{d}})\sqrt{k-1} + O(\frac{1}{k})\sqrt{k-1} \tag{3.20}$$

where $A(.)$ is a function which depends only on x_i and its expression can be found in Mack (1981). Then from assumption 3.9, B_2 vanishes to zero when $n \to \infty$. It remains on showing that B_1 goes to zero also. This can be achieved using result of Mukerjee (1993) or Cheng (1984).

We remark that when the number of neighbors k increases more the weights given to neighbors decrease, then rewriting $m_n(x_i)$ and we have the following equivalence:

$$m_n(x_i) = \frac{\sum_{j=1}^{n} K(\frac{x_i - X_j}{R_i})Y_j}{\sum_{j=1}^{n} K(\frac{x_i - X_j}{R_i})} \cong \sum_{i=1}^{n} \frac{c_j}{k} Y_j \tag{3.21}$$

where $K(.)$ is a given weight function which satisfies condition (3.11), c_j is a bounded weight equal to zero when j larger than the number of neighbors and R_i is the distance between x_i and its k^{th} neighbor. When we denote by $\tilde{m}_n(x_i) = \sum_{i=1}^{n} \frac{c_j}{\sqrt{k}} Y_j$, then from theorem 2.1 in Mukerjee (1993), we have:

$$\begin{aligned} B_1 &= Sup_{x_i} |\tilde{m}_n(x_i) - E[\tilde{m}_n(x_i)]| \\ &= O(\frac{1}{\theta_n}) + O(n^{-\frac{a-1}{a}}) \end{aligned} \tag{3.22}$$

with $a > 1$ and θ_n a positive sequence which tends to zero as $n \to \infty$. So we get $|A_2|$ converges to zero in probability as B_1. This completes the proof of (a).

We have just established asymptotic normality of the encompassing statistic for k-NN regression under the $i.i.d$ assumption. Next, we will study possible extension of such encompassing test to dependent processes.

b2) Dependent case

We study the encompassing test when the processes are dependent. Let reintroduce first the following assumptions on the process and on the regularity condition for the density functions.

Assumption 3.10. *$(S_n)_n$ is ϕ-mixing process.*

Assumption 3.11. *$g(z)$, $\varphi(y \mid z)$ and $\varphi(z)$ are p continuously differentiable. Moreover suppose that $\varphi(y \mid z)$ is bounded.*

Assumption 3.12. *The sequence $k(n) < n$ and such that $\sum_{t=1}^{k(n)} w_t = 1$ where w_i a weight function satisfying $0 < w_t < 1$ when $t \leq k(n)$ and $w_t = 0$ otherwise.*

Under assumptions 3.10-3.12, we have an asymptotic normality of the centered k-NN regression estimate g_n of g, Guégan and Rakotomarolahy (2010). In addition to these assumptions, we make the following assumption to ensure that the bias will be asymptotically negligible.

Assumption 3.13. *$log(n) = o(k)$ and $k = n^\beta$, $0 < \beta < \frac{2p}{2p+d}$.*

We now provide asymptotic normality of the encompassing statistic of functional parameter for time series.

Theorem 3.5. *Suppose that assumptions 3.10-3.13 hold, then under $\mathcal{H}$, we have:*

$$\sqrt{k}\hat{\delta}_{m,g}(z) \rightarrow N(0, \gamma.\sigma^2) \quad in \quad distribution \quad as \quad n \rightarrow \infty. \tag{3.23}$$

where γ a positive constant which is equal to 1 when one considers uniform weights.

Proof of theorem 3.5. We have the following encompassing statistic when replacing the estimator of g_n and $\hat{G}(m_n)$ at a given point z:

$$\begin{aligned}
\sqrt{k}\hat{\delta}_{m_n,g_n}(z) &= \sqrt{k}(\sum_{t=1}^{n} w(z - Z_t)Y_t) - \sqrt{k}(\sum_{t=1}^{n} w(z - Z_t)m_n(x_t)) \\
&= \sqrt{k}(\sum_{t=1}^{n} w(z - Z_t)(Y_t - m(x_t)) + \sqrt{k}(\sum_{t=1}^{n} w(z - Z_t)(m(x_t) - m_n(x_t)) \\
&= C + D
\end{aligned} \tag{3.24}$$

where C is the first expression in RHS of the equality which is a k-NN regression of $\epsilon_t = Y_t - m(x_t)$ on Z_t scaled by the coefficient $\sqrt{k}$. Under $\mathcal{H}$ and with the assumptions of theorem 3.5, theorem **??** (theorem 1 in Guégan and Rakotomarolahy (2010)) provides:

$$C \rightarrow N(0, \gamma.\sigma^2) \quad in \quad distribution \quad as \quad n \rightarrow \infty. \tag{3.25}$$

For the second expression D, we can bound it by taking its supremum with respect to x_t, so we get:

$$\begin{aligned}
|D| &\leq Sup_{x_t}\sqrt{k}|m_n(x_t) - m(x_t)| \\
&\leq Sup_{x_t}\sqrt{k}|m_n(x_t) - E[m_n(x_t)]| + Sup_{x_t}\sqrt{k}|E[m_n(x_t)] - m(x_t)| \\
&= D1 + D2
\end{aligned} \tag{3.26}$$

Using the expression of the bias, in theorem **??** (lemma 1 in Guégan and Rakotomarolahy (2010)), D_2 becomes:

$$D_2 = (Sup_{x_t}A(x_t))O((\frac{k}{n})^{\frac{(1-\beta)p}{d}})\sqrt{k} \tag{3.27}$$

which vanishes to zero when assumption 3.13 holds. $A(.)$ is a function which depends only on x_t and its expression can be found in Guégan and Rakotomarolahy (2010).

Concerning D_1, we can obtain the following inequality:

$$\begin{aligned} D_1 &\leq Sup_{x_t}\sqrt{k}|m_n(x_t) - m_n(x)| + \sqrt{k}|m_n(x) - E[m_n(x)]| + Sup_{x_t}\sqrt{k}|E[m_n(x)] - E[m_n(x_t)]| \\ &= E1 + E2 + E3 \end{aligned} \tag{3.28}$$

If we denote by r_t the distance between x_t and its k_t^{th} neighbors, then the two nearest neighbor estimates $m_n(x_t)$ and $m_n(x)$ are given by:

$$m_n(x) = \sum_{s=1}^{n} w(x - X_s)Y_s \quad and \quad m_n(x_t) = \sum_{s=1}^{n} w(x_t - X_s)Y_s \tag{3.29}$$

where $w(x - X_s) = 0$ when $||x - X_s|| > r_n$ for all $s = 1, ..., n$ and $w(x_t - X_s) = 0$ when $||x_t - X_s|| > r_t$ or $s \geq t$. So, using relation (3.29), we have:

$$E_1 = Sup_{x_t}\sqrt{k}|\sum_{s=1}^{n}(w(x - X_s) - w(x_t - X_s))Y_s| \leq Sup_{x_t}\sum_{s=1}^{n}\frac{c_k}{\sqrt{k}}|Y_s| \tag{3.30}$$

where the last inequality comes from $Sup_{x_t}|(w(x - X_s) - w(x_t - X_s))| \leq \frac{c_k}{k}$ for some positive number c_k. If $Sup_{x_t}\sum_{s=1}^{\infty}||Y_s|| < \infty$, E_1 is negligible. Else $Sup_{x_t}\sum_{s=1}^{\infty}|Y_s| = \infty$ which could be $Sup_{x_t}\sum_{s=1}^{n}|Y_s| \approx M\sqrt{log(n)}$ around infinity for a given positive real M, then $E_1 < M\sqrt{\frac{log(n)}{k}}$ which goes to zero as n goes to infinity under assumption 3.13.

Concerning E_2, for a given non zero real sequences a_{tn}, we have:

$$E_2 = \sqrt{k}|\sum_{t=1}^{n}(w(x - X_t)Y_t - Ew(x - X_t)Y_t)| = \sqrt{k}|\sum_{t=1}^{n}a_{tn}\psi_t| \tag{3.31}$$

where $\psi_t = \frac{1}{a_{tn}}(w(x - X_t)Y_t - Ew(x - X_t)Y_t)$. Applying Tchebyshev inequality on E_2, for $\epsilon > 0$ we have:

$$P(|E_2| > \epsilon) = P(\sqrt{k}|\sum_{t=1}^{n}a_{tn}\psi_t| > \epsilon) \leq \frac{k}{\epsilon^2}E(\sum_{t=1}^{n}a_{tn}\psi_t)^2 \tag{3.32}$$

We know that $E\psi_t = 0$ and ψ_t is a ϕ-mixing process, then from inequality of Yoshihara (1978) relation (3.32) becomes: $P(E_1 > \epsilon) \leq \frac{k}{\epsilon^2}cA_n$ with c is a constant and $A_n = \sum_{t=1}^{n}a_{tn}^2$ where we choose a_{tn} according to $\frac{k}{\epsilon^2}cA_n \to 0$ as $n \to \infty$, then $E_2 \to 0$ in probability. For the last expression E_3, using relation (3.29), we obtain the following result:

$$E_3 = Sup_{x_t}\sqrt{k}|E[m_n(x)] - E[m_n(x_t)]| = Sup_{x_t}\sqrt{k}|\sum_{s=1}^{n}E[(w(x - X_s) - w(x_t - X_s))Y_s]| \tag{3.33}$$

Using iterated conditional expectation, we have:

$$\begin{aligned} E_3 &= Sup_{x_t}\sqrt{k}|\sum_{s=1}^{n} E[(w(x-X_s)-w(x_t-X_s))E[Y_s \mid X_s]]| \\ &= Sup_{x_t}\sqrt{k}|\sum_{s=1}^{n} E[(w(x-X_s)-w(x_t-X_s))m(X_s)]| \end{aligned} \tag{3.34}$$

One can write $m(X_s) = m(X_s) - m(x) + m(x)$ and then (3.34) becomes:

$$\begin{aligned} E_3 &= Sup_{x_t}\sqrt{k}|\sum_{s=1}^{n} E[(w(x-X_s)-w(x_t-X_s))(m(X_s)-m(x)+m(x))]| \\ &= Sup_{x_t}\sqrt{k}|\sum_{s=1}^{n} E[(w(x-X_s)-w(x_t-X_s))(m(X_s)-m(x))]| \end{aligned} \tag{3.35}$$

When $m(.)$ satisfies the Lipschitz condition, then relation (3.35) gives the following inequality:

$$\begin{aligned} E_3 &\leq Sup_{x_t}(c\sqrt{k}\sum_{s=1}^{n} E|w(x-X_s)-w(x_t-X_s)|||X_s-x||) \\ &\leq cr\sqrt{k}Sup_{x_t}(\sum_{s=1}^{n} E|w(x-X_s)-w(x_t-X_s)|) \end{aligned} \tag{3.36}$$

then under assumption 3.13 the RHS of the inequality (3.36) vanishes since $Sup_{x_t}(\sum_{s=1}^{n} E|w(x-X_s)-w(x_t-X_s)|) < \infty$. So E_3 goes to zero as $n \to \infty$. Thus, the proof is established.

Next, we will consider the mixed situation where the owner model has parametric specification and the rival is from nonparametric method.

3.4 Parametric modelling vs nonparametric modelling

In this section, we consider the case that model $\mathcal{M}_1$ is a linear parametric model and $\mathcal{M}_2$ is estimated by nonparametric techniques. Therefore, through out of this section, the hypothesis $\mathcal{H}$ will have linear parametric specification. The encompassing statistic associated to the null $\mathcal{M}_1\mathcal{E}\mathcal{M}_2$ can be written as follows:

$$\hat{\delta}_{\beta,g}(z) = g_n(z) - \hat{G}_L(\hat{\beta})(z), \tag{3.37}$$

where $\hat{G}_L(\hat{\beta})$ is an estimate of the pseudo-true value $G_L(\beta)(z)$ associated with g_n on $\mathcal{H}$, which is defined by $G_L(\beta)(z) = \beta' E[X \mid Z = z]$.

For the nonparametric specification of $\mathcal{M}_2$, we will consider again the two nonparametric methods: the kernel regression method and k-NN regression method.

i) Encompassing test when g is estimated using a kernel regression estimate

We consider the estimate g_n which is the kernel regression estimate of g given in (2.1). Since the rival model g is estimated using kernel method, the various assumptions on kernel density and window width for both independent and dependent case in previous section will be maintained. Moreover, we mention that the unknown conditional means in the hypothesis $\mathcal{H}$ have linear projection specification. We separate the two cases: independent and the dependent variables.

i1) Independent case

The asymptotic property of the encompassing statistic defined in relation (3.37), where the rival model is estimated using kernel regression method, under the i.i.d assumption has been developed in Bontemps et al. (2008). We summarize the asymptotic result in the following theorem.

Theorem 3.6. *Assume that the relation (2.4) is satisfied. When the window width h_{2n} satisfies the usual kernel regularity condition and the assumption (3.1), then under $\mathcal{H}$ with linear specification, we get:*

$$\sqrt{nh_{2n}^q}\hat{\delta}_{\beta,g}(z) \to N(0, \frac{\sigma^2 \int K_2^2(u)du}{\varphi(z)}) \quad in \quad distribution \quad as \quad n \to \infty. \tag{3.38}$$

where $\Omega = Var(Z)^{-1}E[Var(Z \mid X)]Var(Z)^{-1}$ and $\varphi(z)$ is the marginal density of the z_i at z.

For the proof of this theorem, we refer to Bontemps et al. (2008).

We now relax the independent hypothesis which seems to be too restricted in time series modelling.

i2) Dependent case

The process does not need to match the independent assumption. It may exhibit some dependence. In that case, we still can establish the asymptotic normality of the encompassing statistic defined in relation (3.37), using similar assumption as in previous section for dependent processes.

Theorem 3.7. *Assume that relation 2.4 and assumptions 3.3-3.5 are satisfied. Then, under $\mathcal{H}$*

with linear specification and when the bandwidth h_{2n} satisfy kernel regularity condition, we get:

$$\sqrt{nh_{2n}^q}\hat{\delta}_{\beta,g}(z) \to N(0, \frac{\sigma^2 \int K_2^2(u)du}{\varphi(z)}) \quad in \quad distribution \quad as \quad n \to \infty. \tag{3.39}$$

$\varphi(z)$ is the marginal density of the Z at z.

Proof of theorem 3.7 The proof of this theorem will be based on the decomposition of the expression of the encompassing statistic into two parts like in the previous case. Using such techniques, we can write the encompassing statistic as follows:

$$\begin{aligned}
\sqrt{nh_{2n}^q}\hat{\delta}_{\beta,g}(z) &= \sqrt{nh_{2n}^q}(g_n(z) - \hat{G}_L(\hat{\beta})(z)) \\
&= \sqrt{nh_{2n}^q}(\sum_{t=1}^n \frac{K_2(\frac{z-Z_t}{h_n})}{\sum_{t=1}^n K_2(\frac{z-Z_t}{h_{2n}})} Y_t - \sum_{t=1}^n \frac{K(\frac{z-Z_t}{h_n})}{\sum_{t=1}^n K(\frac{z-Z_t}{h_n})} \hat{\beta}' X_t) \\
&= \sqrt{nh_{2n}^q} \sum_{t=1}^n \frac{K_2(\frac{z-Z_t}{h_n})}{\sum_{t=1}^n K_2(\frac{z-Z_t}{h_{2n}})} (Y_t - \beta' X_t) \qquad (3.40) \\
&+ \sqrt{nh_{2n}^q} \sum_{t=1}^n \frac{K(\frac{z-Z_t}{h_n})}{\sum_{t=1}^n K(\frac{z-Z_t}{h_n})} X_t'(\beta - \hat{\beta}) \\
&= D_1 + D_2.
\end{aligned}$$

When assumptions 3.3-3.5 hold, then under $\mathcal{H}$, D_1 converges in distribution to a normal law with mean zero and variance $\frac{\sigma^2}{\varphi(z)} \int K_2^2(u)du$, see Bosq (1998). Concerning D_2, we know that under mixing conditions, the normality asymptotic of the linear process has been established, see Peligrad and Utev (1997, 2006). Therefore, this implies the normality asymptotic of $\sqrt{n}(\beta - \hat{\beta})$. The remaining expression in D_2 vanishes to zero as n tends to infinity. Thus, D_2 converges in distribution to zero. This completes the proof. Next, we consider the k-NN regression estimate for the model $\mathcal{M}_2$.

ii) Encompassing test when g is estimated using k-NN regression estimate

In this section, we assume that we estimate the rival model $\mathcal{M}_2$ using k-NN regression method where the owner model $\mathcal{M}_1$ is still with linear parametric specification. The following theorem provides the asymptotic behavior of the encompassing statistic introduced in relation (3.37) for both independent and dependent processes. We can use kernel method or k-NN regression estimates.

Theorem 3.8. *Assume that the relation (2.4) is sastified. When one of the following points holds:*
(A1) Assumptions 3.6-3.9 and the relation (3.11) are satisfied,
(A2) Assumptions 3.10-3.13 are satisfied.
Then under $\mathcal{H}$, we get:

$$\sqrt{k-1}\hat{\delta}_{\beta,g}(z) \to N(0,\Sigma) \quad in \quad distribution \quad as \quad n \to \infty. \tag{3.41}$$

where for (A1) we have $\Sigma = c\sigma^2 \int w^2(u)du$ with c is the volume of unit ball in $\mathbb{R}^q$ and for (A2) we have $\Sigma = \gamma\sigma^2$ with γ a positive constant which is equal to 1 when one considers uniform weights.

We have provided the asymptotic results of the encompassing statistic respectively for independent and dependent processes where (A1) corresponds to the independent case and (A2) for the dependent case. We now provide the proof of this theorem.

Proof of Theorem 3.8. We now prove the case that the owner model $\mathcal{M}_1$ is the linear regression parametric and the rival model $\mathcal{M}_2$ is the k-NN regression nonparametric. We write the encompassing statistic as follows:

$$\begin{aligned}
\sqrt{k-1}\hat{\delta}_{\beta,g}(z) &= \sqrt{k-1}(g_n(z) - \hat{G}_L(\hat{\beta})(z)) \\
&= \sqrt{k-1}(\sum_{i=1}^{n} W(\frac{z-Z_i}{R_n})Y_i - \sum_{i=1}^{n} \tilde{W}(\frac{z-Z_i}{\tilde{R}_n})\hat{\beta}'X_i) \\
&= \sqrt{k-1}\sum_{i=1}^{n} W(\frac{z-Z_i}{R_n})(Y_i - \beta'X_i) + \sqrt{k-1}\sum_{i=1}^{n} \tilde{W}(\frac{z-Z_i}{\tilde{R}_n})X_i'(\beta - \hat{\beta}) \\
&= N_1 + N_2.
\end{aligned} \tag{3.42}$$

We mention that this equality 3.42 holds since the difference $\sqrt{k-1}\sum_{i=1}^{n} W(\frac{z-Z_i}{R_n})\beta' X_i - \sqrt{k-1}\sum_{i=1}^{n} \tilde{W}(\frac{z-Z_i}{\tilde{R}_n})\beta' X_i$ is negligible.
For the first expression $N_1 = \sqrt{k-1}\sum_{i=1}^{n} W(\frac{z-Z_i}{R_n})\epsilon_i$, with $\epsilon_i = Y_i - \beta' X_i$. When (A1) in theorem 3.8 holds, then using result of Mack (1981) we have:

$$N_1 \to N(0,\Sigma) \quad in \quad distribution \quad as \quad n \to \infty. \tag{3.43}$$

where $\Sigma = c.\sigma^2 \int w^2(u)du$. We have similar asymptotic convergence when (A2) in theorem 3.8 but asymptotic variance $\Sigma = \gamma\sigma^2$, from the asymptotic result in Guégan and Rakotomarolahy

(2010).

For $N_2 = (\beta - \hat{\beta})'\sqrt{k-1}\sum_{i=1}^{n}\tilde{W}(\frac{z-Z_i}{\tilde{R}_n})X_i$, under (A1), we know that the estimate $\sqrt{n}(\beta - \hat{\beta})$ converges in distribution to a normal law Z with mean zero and variance . The remaining part of N_2, the other expression $\frac{\sqrt{k-1}}{\sqrt{n}}\sum_{i=1}^{n}\tilde{W}(\frac{z-Z_i}{\tilde{R}_n})X_i$ converges in distribution to zero. Thus, from Slutsky's theorem, N_2 tends to zero in distribution. Same result can be achived under (A2) where this part vanishes from the uniform convergence in Guégan and Rakotomarolahy (2010). Thus, the proof of theorem 3.8 is established.

We now consider the last case where the owner model $\mathcal{M}_1$ is a nonparametric method and the rival model $\mathcal{M}_2$ is a linear parametric model.

3.5 Nonparametric modelling vs parametric modelling

In this section, we consider the owner model $\mathcal{M}_1$ to be estimated using a nonparametric method and the rival model $\mathcal{M}_2$ to be a linear parametric method. Therefore, the encompassing statistic associated to the null $\mathcal{M}_1\mathcal{E}\mathcal{M}_2$ is given by:

$$\hat{\delta}_{m,\gamma} = \hat{\gamma} - \hat{\gamma}(m_n), \tag{3.44}$$

where $\hat{\gamma}(m_n)$ is an estimate of the pseudo-true value $\gamma(m)$ associated with $\hat{\gamma}$ on $\mathcal{H}$, which is defined by $\gamma(m) = (E[ZZ'])^{-1}E[Zm]$. We estimate the unknown conditional mean m associated to the model $\mathcal{M}_1$ using first the nonparametric kernel regression estimate and next the nonparametric k-NN regression estimate.

***) Encompassing test when m is estimated using kernel regression estimate**

When the estimate of model $\mathcal{M}_1$ is obtained from the kernel regression and the model $\mathcal{M}_2$ is from linear parametric modelling, we establish asymptotic normality of the encompassing statistic introduced in relation (3.44). For the independent case we need the following assumption. The assumption concerns the speed of convergence of the window width h_{1n}.

Assumption 3.14. $\sqrt{n}max(\frac{log(n)}{nh_{1n}^d}, h_{1n}^{2d}) \to 0$ *as* $n \to \infty$.

For dependent case, we will reconsider the assumptions previously done for the use of kernel method. We summarize the asymptotic results in the following theorem.

Theorem 3.9. *Assume that relation (2.4) is satisfied. When the kernel K_1 and the bandwidth h_{1n} satisfy the usual regularity condition and when we have one of the following points:*
(1) The i.i.d assumption, the regularity conditions in linear regression and the assumption 3.14 are satisfied
(2) Assumption 3.3 holds and the kernel regression estimate m_n and the process $(Y_n, X_n)_n$ satisfy assumptions 3.4 and 3.5.
Then, under $\mathcal{H}$, we get:

$$\sqrt{n}\hat{\delta}_{m,\gamma} \to N(0, \Sigma) \quad in \quad distribution \quad as \quad n \to \infty. \tag{3.45}$$

where $\Sigma = plim_{n\to\infty} Var(\sqrt{n}\hat{\delta}_{m,\gamma})$.

Proof of theorem 3.9 First, when we meet the point (1) in theorem 3.9 i.e the processes are independent, the proof of the theorem 3.9 can be found in Bontemps et al. (2008).

Second, for point (2) of theorem 3.9 which corresponds to the case that the processes are dependent. The proof will be based on the decomposition of the expression of the encompassing statistic into two parts like in the previous proves. More precisely, we split the encompassing statistic $\sqrt{n}\hat{\delta}_{m,\gamma}$ into two parts. The first part yields $F_1 = \sqrt{n}(\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i)^{-1}(\frac{1}{n}\sum_{i=1}^{n} Z_i(Y_i - m(x_i)))$ which gives the asymptotic normality of the theorem Peligrad and Utev (1997). The second part is $F_2 = \sqrt{n}(\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i)^{-1}(\frac{1}{n}\sum_{i=1}^{n} Z_i(m(x_i) - m_n(x_i)))$. Again, we bound this by taking the supremum with respect to x_i. Thus, F_2 vanishes to zero from the uniform convergence of $m_n(x_i)$, Bosq (1998). This completes the proof of theorem 3.9. Next, we consider the fact that model $\mathcal{M}_1$ will be estimated using k-NN regression method.

****) Encompassing test when m_n is from k-NN regression estimate**

We suppose that the m_n is a k-NN regression estimate. We state in the following theorem the asymptotic normality of the encompassing statistic in relation (3.44) for both independent and dependent processes. We precise that we use the assumptions introduced in previous section for k-NN regression estimate m_n.

Theorem 3.10. *Assume that relation 2.4 is satisfied. Moreover, assume that:*
(1) the i.i.d assumption, assumptions 3.6-3.9, relation (3.11) and the regularity conditions in linear regression are satisfied.
or
(2) Assumption 3.10 holds and the k-NN estimate m_n and the process $(Y_n, X_n)_n$ satisfy assumptions 3.11-3.13.
Then under $\mathcal{H}$, we get:

$$\sqrt{n}\hat{\delta}_{m,\gamma} \to N(0, \Omega) \quad in \quad distribution \quad as \quad n \to \infty. \tag{3.46}$$

where $\Omega = plim_{n\to\infty} Var(\sqrt{n}\hat{\delta}_{m,\gamma})$, in particular for independent case $\Omega = \sigma^2(E[Z'Z])^{-1}$.

Proof of Theorem 3.10. We mention that the functional parameters m_n is from k-NN regression estimate. We rewrite the associated encompassing statistic as follows:

$$\begin{aligned}
\sqrt{n}\hat{\delta}_{m,\gamma} &= \sqrt{n}(\hat{\gamma} - \hat{\gamma}(m_n)) \\
&= \sqrt{n}((\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i)^{-1}(\frac{1}{n}\sum_{i=1}^{n} Z_i Y_i) - (\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i)^{-1}(\frac{1}{n}\sum_{i=1}^{n} Z_i m_n(x_i))) \\
&= \sqrt{n}(\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i)^{-1}(\frac{1}{n}\sum_{i=1}^{n} Z_i(Y_i - m(x_i))) \\
&+ \sqrt{n}(\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i)^{-1}(\frac{1}{n}\sum_{i=1}^{n} Z_i(m(x_i) - m_n(x_i))) \\
&= L_1 + L_2.
\end{aligned} \tag{3.47}$$

where L_1 corresponds to the first expression in the RHS of the equality (3.47). It coincides to the linear regression of the error ϵ on Z, with $\epsilon_i = Y_i - m(x_i)$.
When the processes satisfy independent condition, point (1) in theorem 3.10, then L_1 converges in distribution to Z where Z is normally distributed with mean zero and variance $\Omega = \sigma^2(E[Z'Z])^{-1}$. For the second expression L_2, we bound it by taking the maximum with respect to x_i and then we get $|L_2| \leq \sqrt{n}S_n D_n Sup\{(m(x_i) - m_n(x_i))), x_i \in \mathbb{R}^d\}$ where $S_n = \sum_{i=1}^{n} |Z_i|$ and $D_n = (\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i)^{-1}$. We remark that $\sqrt{n}S_n$ asymptotically converges to a normal distribution with mean $E[|Z|]$ and variance $Var(|Z|)$. Moreover, the remaining expression involves the supremum which vanishes to zero. Thus, the product vanishes to zero also from Slutsky's theorem.
When the processes are dependent, point (2) in theorem 3.10, then the normality asymptotic of

the first expression L_1 is achieved using the normality asymptotic result for dependent processes in Peligrad and Utev (1997). Similarly, the second expression L_2 vanishes from the same asymptotic normality of linear process in Peligrad and Utev (1997). This completes the proof.

We now mention some points about the encompassing test as well as about its asymptotic behaviors developed in previous paragraphs. For both independent and dependent cases, we have some remarks concerning the hypotheses on the encompassing test as well as on the asymptotic results of the encompassing statistics.

Remark 3.1. *The two hypothesis $\mathcal{H}$ and $\mathcal{H}^*$ are not at all equivalent. As pointed in Bontemps et al. (2008) the implication $\mathcal{H} \Rightarrow \mathcal{H}^*$ always holds and equivalence is proved only for linear models.*

Remark 3.2. *We should be careful also about mutual encompassing of both models which concerns the bijection of the pseudo true value function $G(.)$.*

Remark 3.3. *We have considered the same number of neighbors for different regressions in encompassing test associated to nearest neighbor regression. In general, we have two different number of neighbors entering in the proof of theorems which correspond to the k-NN estimate of m and g, but the asymptotic results remain valid.*

In the following remarks, we take a view on the rate of convergences and asymptotic variances of the asymptotic encompassing statistic to k-NN and kernel regression estimates.

Remark 3.4. *When we take a look at the convergence rates of the asymptotic encompassing statistic associated to k-NN and kernel regressions, it depends only on the number of neighbors k for k-NN while for kernel it depends on the number of observation n and the bandwidth h_n. We have the same convergence rate when $h_n = k/n$.*

Remark 3.5. *Concerning the variances of the asymptotic encompassing statistic for both regression methods, we see that the asymptotic variance of the encompassing statistic associated to kernel regression depends on the density, which is not the case for nearest neighbor regression estimate.*

These last two remarks appear in both independent and dependent processes. They show some advantages of k-NN nearest neighbor regression. We now provide an application on real data of the asymptotic result on encompassing test.

4 Illustration

In this section, we provide example on possible application of the above theoretical results. We focus on economic modelling and consider the problem about the relationship between financial variables and real economic activity such as modelling the impact of the interest rate changes on the Gross Domestic Product (GDP). We consider five variables which are well known to affect the behavior of the GDP. We will use the linear autoregressive modelling and the nearest neighbor nonparametric regression method. The last modelling is interesting as it is known to take nonlinearity into account.

We use quarterly data of US real GDP from 1995Q1 to 2010Q1 and the following US monthly data from 1995M1 to 2010M3: stock market index (Index), real estate price index, oil price index, real effective exchange rate (EER) and interest rate spread. The last variable is obtained by taking the difference between ten year and three month government bonds. In general, we transform the monthly variables into quarterly by aggregating the monthly values in a quarter.

For the linear modelling, we look at the dynamic of the centered logarithm GDP. From the stationary analysis, we consider a linear autoregressive representation with determistic trend. We use several criteria for model selection. They are the out-of-sample selection FPE and the two information criteria BIC and the corrected AIC. We consider also the apriori order four that we often meet in autoregressive modelling of the quarterly GDP. Then using sample from 1995Q1 to 2005Q1, AICc and BIC criteria result to AR(1) denoted by M_1 and the out-of-sample FPE criterion yields AR(3) denoted by M_2. The third model AR(4) is denoted by M_3 . The decision on choosing one model will be based on encompassing test. A necessary condition is that the encompassing model should fit better than encompassed model. Therefore, encompassing model is expected to have smaller error variance than its rival. The standard errors of models M_i, $i = 1, 2, 3$ are $\sigma_1 = 0.0050$, $\sigma_2 = 0.0047$ and $\sigma_3 = 0.0048$, respectively. Then, among the three models, M_1 has the worst fit. In contrast, M_2 has the best fit. For the encompassing test between two linear parametric models, we present the associated results in table 1.

From table 1, we accept the null $M_2 \mathcal{E} M_3$ that is, M_2 encompasses M_3. In contrast, we reject

Table 1: Encompassing test for the logarithm GDP model.

	$M_1\mathcal{E}M_3$	$M_3\mathcal{E}M_1$	$M_3\mathcal{E}M_2$	$M_2\mathcal{E}M_3$
t_{encomp}	0.79 (0.43)	11.50 (0.00)	3.95 (0.00)	1.03 (0.31)

Notes: Statistics of the test with their p-values in parenthesis.

$M_3\mathcal{E}M_2$ and $M_1\mathcal{E}M_2$ and therefore M_3 encompasses neither M_1 nor M_2. Moreover, M_2 has the smallest error variance. Thus, we retain model M_2 which is from the out-of-sample selection criterion.

We now consider the nonparametric nearest neighbor regression method. We will use the following nonlinear representation

$$y_{t+1} = \beta' z_t + m_y(z_t) + u_t \quad and \quad x_{t+1} = m_x(x_t) + v_t \tag{4.1}$$

where $z_t = (y_t, x_t)'$ with y_t is the logarithm GDP, x_t is the logarithm of financial/economic variables except for the interest rate spread which is always in level, u_t and v_t are error processes and the uknown parameters are β and the two nonlinear functions m_y and m_x.

We use the same information set for the estimation sample as in previous modelling. For the estimation procedure, we estimate the parameters in two steps where we first regress y_{t+1} on z_t for the parameter β, we next recuperate the estimated residuals. We then use the estimated residuals for the estimation of functions m_y and m_x using nearest neighbor method with a priori choice for the number of neighbors $k = 3$ and for the dimension $d = 2$. We now have model from estimation step. Next, we are interested on testing if the nearest neighbor regression technique can encompass linear model. We apply the encompassing test on nonparametric and parametric regression techniques developed previously having as null hypothesis: the early nearest neighbor regression of GDP with exogenous variables encompasses the linear AR(3) for the GDP. Under this null, we have the statistic $S = \frac{\hat{\delta}}{\sqrt{\hat{\Omega}}}$, which is approximately standard normal distribution, where $\hat{\Omega}$ is an estimate of the asymptotic variance Ω. That is we have $S = \frac{\hat{\delta}}{\sqrt{\hat{\Omega}}} = (\sum_t Z_t Z_t)^{-1} \sum_t u_t Z_t / \sqrt{\hat{\Omega}}$ where $\hat{\Omega} = \hat{\sigma}^2 (\sum_t Z_t Z_t)^{-1}$, with $Z_t = y_{t-3}$ is the regressor of the model AR(3), $\hat{u}_t$ is the estimated residuals from nearest neighbor model and $\hat{\sigma}^2$ is a k-NN estimate of the conditional variance $\sigma^2 = var(y_t/y_{t-1}, y_{t-3}, x_{t-1})$, recalling that y_t is the centered logarithm GDP and x_t the logarithm financial/economic variables. We report the result in table 2. We remark that,

Table 2: Encompassing test nearest neighbor vs autoregressive.

test	Statistic for horizon h (p-value in parenthesis)			
kNN$\mathcal{E}$AR(3)	1	2	4	8
Spread$\mathcal{E}$AR(3)	-0.75 (0.45)	-1.32 (0.20)	-1.83 (0.08)	-0.55 (0.58)
Oil$\mathcal{E}$AR(3)	-0.67 (0.50)	-1.00 (0.32)	-0.91 (0.37)	-0.22 (0.22)
Housing$\mathcal{E}$AR(3)	-0.45 (0.65)	-0.91 (0.37)	-0.47 (0.63)	0.04 (0.96)
Index$\mathcal{E}$AR(3)	-0.36 (0.71)	-0.72 (0.47)	-0.59 (0.56)	0.01 (0.99)
EER$\mathcal{E}$AR(3)	-0.48 (0.63)	-0.88 (0.38)	-0.56 (0.57)	-0.06 (0.94)

given a risk level for example five percent, overall we accept the null. This implies that k-NN modelling of GDP with exogenous variables encompasses the model AR(3) of GDP. In other words, information content on the model AR(3) is already included in the k-NN modelling of GDP with financial/economic variables. Such failure of the model based on the sole GDP to encompass the model based on the five variables highlights that economic and financial variables contain information unexplained in the GDP dynamic. We remark the presence of nonlinearity on the relationship of economic/financial variables with GDP. Accounting such nonlinear feature on the model would be essential when assessing the role of those economic and financial variables on forecasting GDP.

5 Conclusion

We know that different approaches of encompassing tests present in the literature provide different results. We have considered encompassing test in asymptotic way which is inline with the encompassing principle announced in the introduction. The work has been conducted for nonparametric methods.

As stated in Hendry et al. (2008) that the work of Bontemps *et al.* (2008) is the starting treatment of encompassing tests to functional parameter based on nonparametric methods. We have extended that work first to nearest neighbor functional parameter estimate under the i.i.d assumption and second to dependent process.

When using nearest neighbor regression as estimator for conditional expectations, we have established asymptotic normality of the encompassing test for independent processes. Same results have been provided for dependent processes, considering both kernel and nearest neighbor regressions.

Development of encompassing test to nonparametric methods opens new research direction in theory as well as in practice. Application of the various results on real data would accelerate such development.

References

[1] Bontemps, C., J.P. Florens and J.F. Richard (2008), "Parametric and non-parametric encompassing procedures", Oxford Bulletin of Economics and Statistics 70, 751–780.

[2] Bontemps, C. and G.E. Mizon (2008). "Encompassing: concepts and implementation", Oxford Bulletin of Economics and Statistics 70, 721–750.

[3] Bosq, D. (1998), "Nonparametric Statistics for Stochastic Processes: Estimation and Prediction", Lecture Notes in Statistics, Springer-Verlag, 2nd eds, Berlin.

[4] Cheng, P.E. (1984), "Strong consistency of nearest neighbor regression function estimators", Journal of Multivariate Analysis 15, 63-72.

[5] Florens, J.P., D.F. Hendry and J.F. Richard (1996), "Encompassing and specificity", Econometric Theory 12, 620–656.

[6] Gouriéroux, C. and A. Monfort (1995), "Testing, encompassing, and simulating dynamic econometric models", Econometric Theory 11, 195–228.

[7] Gouriéroux, C., A. Monfort and A. Trognon (1983), "Testing nested or non-nested hypotheses", Journal of Econometrics 21, 83–115.

[8] Govaerts, B., D.F. Hendry and J.F. Richard (1994), "Encompassing in stationary linear dynamic models", Journal of Econometrics 63, 245–270.

[9] Guégan, D. and N. Huck (2005), "On the use of nearest neighbors in finance", Revue de Finance 26, 67-86.

[10] Guégan, D. and P. Rakotomarolahy (2009), "The multivariate k-nearest neighbor model for dependent variables : one-sided estimation and forecasting", CES Working Paper, No. 2009.50.

[11] Guégan, D. and P. Rakotomarolahy (2010), "A short note on the nowcasting and the forecasting of Euro-area GDP using non-parametric techniques", Economics Bulletin 30, 508-518.

[12] Guégan, D. and P. Rakotomarolahy, (2010), " Alternative methods for forecasting GDP", Chapter 8 in Fredj Jawadi, William A. Barnett (ed.) Nonlinear Modeling of Economic and Financial Time-Series (International Symposia in Economic Theory and Econometrics, vol. 20), Emerald Group Publishing Limited, 161-185.

[13] Hendry, D. F. (1995), "Dynamic Econometrics", Oxford University Press, Oxford.

[14] Hendry, D.F. and J.A. Doornik (1994), "Modelling linear dynamic econometric systems", Scottish Journal of Political Economy 41, 1-33.

[15] Hendry, D.F., M. Marcellino and G.E. Mizon (2008), "Encompassing", Special Issue: Oxford Bulletin of Economics and Statistics, Guest Editor Introduction.

[16] Hendry, D.F. and B. Nielsen (2006), "Econometric modeling: a Likelihood Approach", Princeton University Press, Princeton.

[17] Hendry, D. F. and J.F. Richard (1989), "Recent developments in the theory of encompassing", in Cornet B. and Tulkens H. (eds), Contributions to Operations Research and Economics, The XXth Anniversary of CORE, 393–440, Cambridge.

[18] Hoover, K.D. and S.J. Perez (1999), "Data mining reconsidered: encompassing and the general-to-specific approach to specification search", Econometrics Journal 2, 167–191.

[19] Mack, Y.P. (1981), "Local properties of k-NN regression estimates", SIAM Journal on Algebraic and Discrete Methods 2, 311-323.

[20] Mizon, G. E. (1984), "The encompassing approach in econometrics", in D.F. Hendry and K.F. Wallis (eds), Econometrics and Quantitative Economics, 135–172.

[21] Mizon, G. E. (2008). "Encompassing", in Blume L. E. and Durlauf S. N. (eds), The New Palgrave Dictionary of Economics, 2nd edn. The New Palgrave Dictionary of Economics Online, Palgrave Macmillan. http://www.dictionaryofeconomics.com/article?id=pde2008 E000243.

[22] Mizon, G. E. and J.F. Richard (1986), "The encompassing principle and its application to non-nested hypothesis tests", Econometrica 54, 657–678.

[23] Mizrach, B. (1992), "Multivariate Nearest-Neighbour Forecasts of EMS Exchange Rates", Journal of Applied Econometrics 7, Supplement: Special Issue on Nonlinear Dynamics and Econometrics (Dec., 1992), S151-S163.

[24] Mukerjee, H. (1993), "Nearest neighbor regression with heavy-tailed errors", Annals of Statistics 21, 681-693.

[25] Nowman, B. and B. Saltoglu (2003), "Continuous time and nonparametric modelling of U.S. interest rate models", International Review of Financial Analysis 12, 25-34.

[26] Peligrad, M. and S.A. Utev (1997), "Central limit theorem for Linear Processes", The Annals of Probability 25, 443-456.

[27] Sawa, T. (1978), "Information criteria for discriminating among alternative regression models", Econometrica 46, 1273–1292.

[28] Vieu, P. (1994), "On variable selection in nonparametric regression", Computation Statistics and Data Analysis 17, 575–594.

[29] Yoshihara, K. (1978). "Probability inequalities for sums of absolutely regular processes and their applications", Zeitschrift fuer Wahrscheinlichkeitstheorie und Verwandte Gebiete 43, 319–329.